

Deep-learning prediction of progression-free survival to EGFR inhibitors from H&E tissue slides in advanced *EGFR*-mutated non-small cell lung cancer

→ **Lodovica Zullo**^{1*}, Elin Samuelsson^{2*}, Fabrizio Citarella³, Mina Farag², Francesco Cortiula⁴, Andrea De Giglio⁵, Frank Aboubakar⁶, Katharina von Loga², Etienne Hatton², Lorenzo Nibid⁷, Giovanna De Maglio⁸, Francesca Ambrosi⁹, Beatrice Ramella Pollone¹⁰, Melanie Janson¹¹, Alessandro Russo¹², Andrew Whittum¹³, Ingrid Garberis¹³, Victor Aubert², Hortense Deslandes², David Planchard¹, Fabrice André¹, Cecile Badoual¹³, Benjamin Besse¹, Maria-Rosa Ghigna^{13*}, Mihaela Aldea^{1*}

¹Université Paris-Saclay, Gustave Roussy, Inserm, Département de Médecine Oncologique, F-94805, Villejuif, France, ²Owkin, Paris, France, ³Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Roma, Italy, ⁴Department of Oncology, University Hospital of Udine, Udine, Italy, ⁵Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy, ⁶Division of Pneumology, Cliniques Universitaires Saint-Luc, Brussels, Belgium, ⁷Research Unit of Anatomical Pathology, Department of Medicine and Surgery, Università Campus Bio-Medico di Roma, Roma, Italy, ⁸Department of Pathology, University Hospital of Udine, Udine, Italy, ⁹Pathology Unit, Maggiore Hospital-AUSL Bologna, Bologna, Italy; Department of Experimental, Diagnostic and Specialty Medicine (DIMES), University of Bologna, Bologna, Italy, ¹⁰Lung Cancer Unit, IRCCS Ospedale Policlinico San Martino, Genova, Italy, ¹¹François BACLESSE, Comprehensive Cancer Center, Caen, France, ¹²Medical Oncology, A.O. Papardo & Department of Human Pathology, University of Messina, Messina, Italy, ¹³Université Paris-Saclay, Gustave Roussy, Inserm, Département de Biologie et Pathologie Médicale, F-94805, Villejuif, France.

Background

Third-generation tyrosine kinase inhibitors (TKIs) like osimertinib are standard treatments for advanced, previously untreated non-small cell lung cancer (NSCLC) with common *EGFR* mutations (exon 19 deletion/L858R). Combination with chemotherapy, bispecific antibodies or anti-angiogenic agents prolongs survival but raises toxicity. To date, no biomarker identifies patients benefiting from combinations versus osimertinib alone.

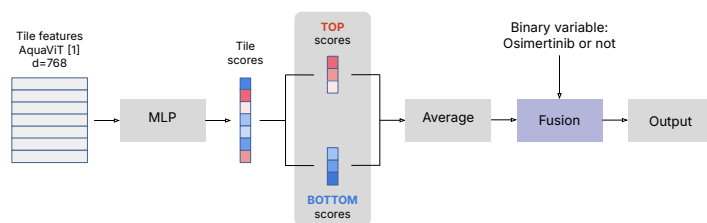
Objective

To develop and validate a multiple instance learning (MIL) model to predict progression-free survival (PFS) under EGFR-TKIs in patients with advanced NSCLC and *EGFR* common mutations, using digitized H&E tissue slides and accounting for the type of administered EGFR-TKI.

Data

This is a multi-center, retrospective study including patients with advanced *EGFR*-mutated NSCLC, treated with first-line TKIs, either osimertinib or first-generation inhibitors (erlotinib/gefitinib). Three independent cohorts were collected from Gustave Roussy (Villejuif, France), Campus Bio-medico (Rome, Italy) and the University Hospital of Udine (Udine, Italy). The model was trained on the Gustave Roussy cohort, while the Campus Bio-Medico and Udine cohorts were reserved for independent external validation.

Methods



Whole-slide digitization (MPP ~0.5) of diagnostic H&E-slides (lung samples) obtained <= 60 days prior to the start of first line treatment, conducted locally at each institution using different scanners.

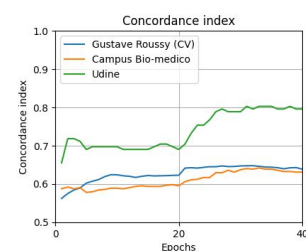
Training step 1: Weakly supervised MIL model Weldon [2].

Training step 2: Fusion = logit + Δ, where the correction term Δ is generated by a single fully connected layer taking as input the logit and a binary variable capturing the type of administered EGFR-TKI.

Predictive Performance

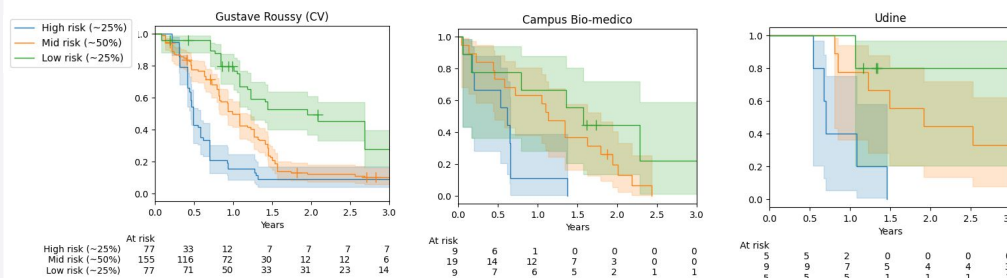
The Gustave Roussy training cohort consists of 309 WSIs from 81 patients, while the external validation cohorts from Campus Bio-medico and Udine consists of one WSI from each of 37 and 27 included patients, respectively. Eight cytology samples were dropped from the Udine cohort, leaving only 19 patients for analysis.

Across the three cohorts (Gustave Roussy, Campus Bio-medico, Udine), patients were mostly female (65%, 62%, 74%) with no smoking habit (62%, 57%, 53%), had adenocarcinoma (93%, 95%, 100%) harboring exon 19 deletions (56%, 54%, 58%). The median (IQR) age at diagnosis was 65 (56-74), 71 (62-79), 68 (60-76) years.



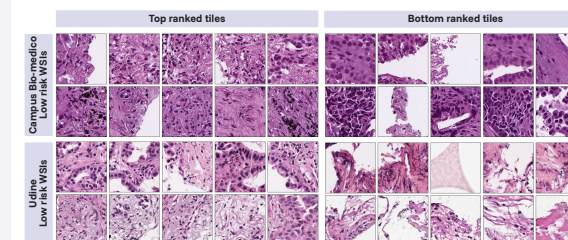
| Cohort | Subset | # Patients | True PFS in months median (95%CI) | C-index for continuous PFS mean (95%CI) | AUROC for 1-year PFS mean (95%CI) |
|-------------------|------------------|------------|-----------------------------------|---|-----------------------------------|
| Gustave Roussy | All | 81 | 12.1 (9.8, 15.4) | Training cohort | |
| | Osimertinib | 42 | 12.6 (9.8, 16.9) | | |
| | First-generation | 39 | 10.1 (7.0, 15.4) | | |
| Campus Bio-medico | All | 37 | 12.6 (7.4, 16.5) | 0.64 (0.53, 0.74) | 0.74 (0.61, 0.86) |
| | Osimertinib | 17 | 16.1 (5.5, 23.3) | 0.68 (0.50, 0.85) | 0.62 (0.35, 0.87) |
| | First-generation | 20 | 7.8 (2.1, 16.5) | 0.53 (0.41, 0.65) | 0.67 (0.45, 0.86) |
| Udine | All | 19 | 17.9 (10.2, 36.4) | 0.79 (0.68, 0.89) | 0.86 (0.71, 1.00) |
| | Osimertinib | 8 | 30.3 (10.2, 50.1) | - | - |
| | First-generation | 11 | 13.0 (8.0, 17.9) | 0.76 (0.58, 0.92) | 0.76 (0.50, 1.00) |

We defined high-, intermediate-, and low-risk patients by splitting at the first and third quartiles of the predicted score per cohort.

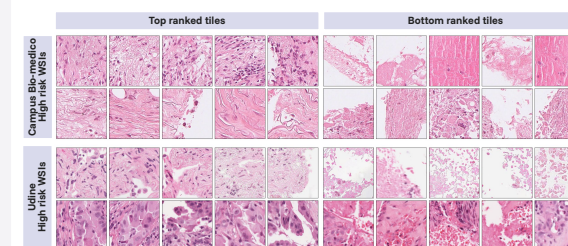


Interpretability

From each validation set, five WSIs with longest and shortest predicted PFS were selected for manual inspection. Their top- and bottom-ranked tiles were reviewed to assess whether the model focuses on meaningful regions and captures biologically relevant patterns. Examples of such extreme tiles are shown below.



Top-ranked tiles in WSIs with long PFS prediction (low risk) predominantly captures tumoral regions with acinar architecture, intermediate (G2) cytological grading, and a fibrosis-rich microenvironment. The bottom-ranked tiles often focus on less informative areas or, more rarely, on hyperchromatic nuclei.



For short PFS prediction (high risk) WSIs, the top-ranked tiles show tumoral cells with different grade of pleomorphism and architectures. The bottom-ranked tiles often display necrotic foci.

Conclusions

We developed a deep learning model to predict PFS to EGFR-TKI from diagnostic H&E slides. The model was trained and externally validated on histological lung specimens, with promising discriminative performance. Calibration remains to be refined in larger cohorts. Further validation is ongoing using samples from the EORTC Lung Cancer Group-1613 phase II randomized trial, including patients treated with first-line osimertinib or gefitinib.

References

- [1] A. Filot et al., "Distilling foundation models for robust and efficient models in digital pathology," MICCAI 2025.
- [2] T. Durand et al., "WELDON: Weakly Supervised Learning of Deep Convolutional Neural Networks," CVPR 2016.